CLAIMS

1. A system of intercepting multimedia documents disseminated from a first network (A), the system being characterized in that it comprises a module (110) for intercepting and processing packets of information each including an identification header and a data body, the packet interception and processing module (110) comprising first means (101) for intercepting packets disseminated from the first network (A), means (102) for analyzing the headers of packets in order to determine whether a packet under analysis forms part of a connection that has already been set up, means (104) for processing packets recognized as forming part of a connection that has already been set up to determine the identifier of each received packet and to access a storage container (115) where the data present in each received packet is saved, and means (103) for creating an automaton for processing the received packet belonging to a new connection if the packet header analyzer means (102) show that a packet under analysis constitutes a request for a new connection, the means (103) for creating an automaton comprise in particular means for creating a new storage container (115) for containing the resources needed for storing and managing the data produced by the means (104) for processing packets associated with the new connection, a triplet comprising <identifier, connection state flag, storage container> being created and being associated with each connection by said means (103) for creating an automaton, and in that it further comprises means (111) for analyzing the content of data stored in the containers (115), for recognizing the protocol used from a set of standard protocols such as in particular http, SMTP, FTP, POP, IMAP, TELNET, P2P, for analyzing the content transported by the protocol, and for reconstituting the intercepted documents.

2. An interception system according to claim 1, characterized in that the analyzer means (102) and the processor means (104) comprise a first table (116) for setting

up a connection and containing for each connection being set up an identifier "connectionId" and a flag "connectionState", and a second table (117) for identifying containers and containing, for each connection that has already been set up, an identifier "connectionId" and a reference "containerRef" identifying the container (115) dedicated to storing the data extracted from the frames of the connection having the identifier "connectionId".

3. An interception system according to claim 2, characterized in that the flag "connectionState" of the first table (116) for setting up connections can take three possible values (P10, P11, P12) depending on whether the detected packet corresponds to a connection request made by a client, to a response made by a server, or to a confirmation made by the client.

4. An interception system according to any one of claims 1 to 3, characterized in that the first packet interception means (101), the packet header analyzer means (102), the automaton creator means (103), the packet processor means (104), and the means (111) for analyzing the content of data stored in the containers (115) operate in independent and asynchronous manner.

5. An interception system according to any one of claims 1 to 4, characterized in that it further comprises a first module (120) for storing the content of documents intercepted by the module (110) for intercepting and processing packets, and a second module (121) for storing information relating to at least the sender and the destination of intercepted documents.

6. An interception system according to claim 5, characterized in that it further comprises a module (122) for storing information relating to the components that result from detecting the content of intercepted documents.

7. An interception system according to any one of claims 1 to 6, characterized in that it further comprises a centralized system (200) comprising means (221) for producing fingerprints of sensitive documents under surveillance, means (222) for producing fingerprints of intercepted documents, means (220) for storing fingerprints produced from sensitive documents under surveillance, means (250) for storing fingerprints produced from intercepted documents, means (260) for comparing fingerprints coming from the means (250) for storing fingerprints produced from intercepted documents with fingerprints coming from the means (220) for storing fingerprints produced from sensitive documents under surveillance, and means (213) for processing alerts, containing the references of intercepted documents that correspond to sensitive documents.

8. An interception system according to claim 7, characterized in that it includes selector means (130) responding to the means (213) for processing alerts to block intercepted documents or to forward them towards a second network B, depending on the results delivered by the means (213) for processing alerts.

9. An interception system according to claim 7 or claim 8, characterized in that the centralized system (200) further comprises means (230) for associating rights with each sensitive document under surveillance rights, and means (240) for storing information relating to said rights, which rights define the conditions under which the document can be used.

10. An interception system according to any one of claims 1 to 9, characterized in that it is interposed between a first network A of the LAN type and a second network B of the LAN type.

11. An interception system according to any one of claims 1 to 9, characterized in that it is interposed between a first

network A of the Internet type and a second network B of the Internet type.

12. An interception system according to any one of claims 1 to 9, characterized in that it is interposed between a first network A of the LAN type and a second network B of the Internet type.

13. An interception system according to any one of claims 1 to 9, characterized in that it is interposed between a first network A of the Internet type and a second network B of the LAN type.

14. An interception system according to claim 13, characterized in that it further comprises a generator (300) for generating requests from sensitive documents to be protected, in order to inject requests into the first network A.

15. An interception system according to claim 14, characterized in that the request generator (300) comprises:

· means (301) for producing requests from sensitive documents under surveillance;

· means (302) for storing the requests produced;

· means (303) for mining the first network A with the help of at least one search engine using the previously stored requests;

· means (304, 306) for storing the references of suspect files coming from the first network A; and

· means (305) for sweeping up suspect files referenced in the means (304, 306) for storing references and for sweeping up files from the neighborhood, if any, of the suspect files.

16. An interception system according to any one of claims 7 to 9, characterized in that said means (260) for comparing fingerprints deliver a list of retained suspect documents (211) having a degree of pertinence relative to sensitive documents, and the alert processor means (213) deliver the

references of an intercepted document when the degree of pertinence of said document is greater than a predetermined threshold.

17. An interception system according to any one of claims 7 to 9, characterized in that it further comprises, between said means (260) for comparing fingerprints and said means (213) for processing alerts, a module (212) for calculating the similarity between documents, which module comprises:

a) means for producing an interference wave representing the result of pairing between a concept vector taken in a given order defining the fingerprint of a sensitive document and a concept vector taken in a given order defining the fingerprint of a suspect intercepted document; and

b) means for producing an interference vector from said interference wave enabling a resemblance score to be determined between the sensitive document and the suspect intercepted document under consideration, the means (213) for processing alerts delivering the references of a suspect intercepted document when the value of the resemblance score for said document is greater than a predetermined threshold.

18. An interception system according to any one of claims 7 to 9, characterized in that it further comprises, between said means (260) for comparing fingerprints and said means (213) for processing alerts, a module (212) for calculating similarity between documents, which module comprises means for producing a correlation vector representative of the degree of correlation between a concept vector taken in a given order defining the fingerprint of a sensitive document and a concept vector taken in a given order defining the fingerprint of a suspect intercepted document, the correlation vector enabling a resemblance score to be determined between the sensitive document and the suspect intercepted document under consideration, the means (213) for processing alerts delivering the references of a suspect intercepted document when the value of the resemblance score for said document is greater than a predetermined threshold.